Integrating Machine Learning with Open-Source GIS for Reproducible High-Resolution Terrain Classification and Hazard Mapping from UAV LiDAR Data

Nikola Kranjčić [1], Olga Bjelotomić Oršulić [1], Hrvoje Matijević [1], Saša Vranić [1]

[1] University North, Department of Geodesy and Geomatics, Jurja Križanića 31b, 42000 Varaždin, Croatia - nkranjcic@unin.hr, oborsulic@unin.hr, hmatijevic@unin.hr, svranic@unin.hr

Keywords: UAV LiDAR, open-source GIS, machine learning, geomorphometry, terrain classification, hazard susceptibility mapping

Abstract

This study provides an approachable, open-source methodology for terrain classification and hazard-susceptibility mapping based on UAV LiDAR, geomorphometry, and machine learning. First, the terrain model is calculated from UAV LiDAR data after ground-point classification, outliers' exclusion, and data interpolation. Geomorphometry variables are then computed based on the terrain model, including slope, aspect, plan and profile curvatures, TPI, terrain roughness, flow accumulation, and LS-factor. Reference data required for supervised classification come from the geomorphological analysis of LiDAR-based terrain model and engineering-geological information. Stable and unstable terrain units are used to build and test Random Forest, SVM, and GB models. Metrics such as the confusion matrix, overall accuracy, precision, recall, F1-score, ROC-AUC allow for evaluating the results of models, and feature importance determines the key predictors. The main novelty lies in an easily reproducible methodology using only open-source software, including QGIS, PDAL, and Python. The introduced methodology makes it possible to perform validation and code reuse, thereby transferring the UAV LiDAR terrain hazard analysis to any area of research in the open geospatial environment.

1. Introduction

High-resolution topographic data provide a crucial tool for terrain morphological analysis, identification of geomorphological processes and terrain hazards evaluation. Modern UAV-based LiDAR technology has significantly improved the possibility of acquiring high-resolution 3D point clouds in site-specific scales. Compared to traditional topographic surveys or low-resolution elevation models, UAV LiDAR allows generating high-resolution representation of terrain surfaces and detecting small scale morphological elements such as slope breaks, scarps, incised drainages, surface roughness and terrain discontinuities caused by erosion (Barbarella et al., 2021; Bartmiński et al., 2023; Storch et al., 2023). All these characteristics make UAV LiDAR applicable for terrain classification and susceptibility assessment in case of slope instability, erosion and complex morphodynamics (Barbarella et al., 2021; Bartmiński et al., 2023; Storch et al., 2023).

Machine learning techniques are becoming increasingly common in terrain hazard and landslide susceptibility studies due to the capability to model non-linear relations between terrain variables and potential hazard occurrence. Random Forest, Support Vector Machine, Gradient Boosting and XGBoost models are actively used in geospatial classifications and showed good predictive performance in susceptibility maps (Kavzoglu and Teke, 2022; Scheiber et al., 2026; Xu et al., 2024). Among these studies, geomorphometric variables derived from digital terrain models are considered as the most influential predictors. Such morphological variables as slope gradient, aspect, plan and profile curvature, terrain position index, topographic roughness, flow accumulation, wetness index, erosion-related indexes are commonly used to characterize terrain shape, hydrological concentration and process potential (Al-Najjar et al., 2021). However, the reliability of these features strongly depends on quality and resolution of the digital terrain models utilized as an input dataset.

Despite wide utilization of machine learning algorithms in susceptibility modelling, there are several weaknesses in current approaches. Some studies focus on maximizing predictive accuracy while neglecting issues of the reproducibility of the full workflow. In particular, data pre-processing, point-cloud filtering, creation of the digital terrain model, raster-derivative calculations, sampling strategy, model training and validation process, output map production may be only partially described. Moreover, some pipelines include proprietary tools which limit their accessibility, transparency and ability to reproduce and use independently by other researchers (Forkan et al., 2023; Maesen and Salingros, 2024; Storch et al., 2025).

However, open-source geospatial tools and programming languages provide a solid basis to tackle these problems. QGIS, GRASS GIS, SAGA GIS, PDAL, ESA SNAP, Python and many others provide comprehensive tools for developing reproducible workflow based on open geospatial standards. Reproducibility becomes not only a technical characteristic but also an important contribution to open science in the field of geomatics and geospatial science (Butler et al., 2021; Forkan et al., 2023; Storch et al., 2025).

In this paper, we propose an open-source and reproducible workflow for high-resolution terrain hazard mapping and susceptibility assessment based on geomorphometric extracted from UAV LiDAR point clouds and machine learning. We present and discuss a pipeline which includes point-cloud filtering, creation of the digital terrain model, generation of geomorphometric features, machine learning and feature-importance analysis. The comparison of Random Forest, Support Vector Machine and Gradient Boosting models was performed using quantitative validation metrics including accuracy, precision, recall, F1-score, ROC-AUC and confusion matrix analysis. The goal was to assess the impact of geomorphometric terrain variables on predicting hazardous areas based on a given dataset.

The novelty of this study consists of three aspects. First, the applicability of an open-source workflow for hazard and susceptibility assessments based on UAV LiDAR is shown. Second, a reproducible machine learning pipeline based on geomorphometric predictions is developed. Third, the interpretability of machine learning models in terms of terrain parameters is highlighted.

2. Study area and data

The research was conducted on a UAV LiDAR survey field located in Croatia. The site is characterized by heterogeneity of terrain morphology with some slope geomorphological processes locally represented there. The studied site can be considered as a suitable location for a case study focused on terrain classification in terms of high-resolution geomorphometric, because it involves stable and potentially unstable terrain units, slopes with a complicated morphology, drainage forms, and areas potentially affected by some geomorphological processes related to the erosion process.

An initial topographic dataset used for geomorphometric feature calculation was produced through a UAV LiDAR survey. The dataset involved three-dimensional point cloud obtained from the survey made by a DJI Matrice 350 RTK drone with a DJI Zenmuse L2 LiDAR sensor. Surveying was performed with RTK corrections obtained from the Croatian Positioning System (CROPOS) therefore the obtained point cloud can be successfully utilized for the construction of digital terrain models and other analyses.

A point cloud from UAV LiDAR was classified, outliers were removed, and a digital terrain model was created from the remaining data. The resulting DTM was generated in the official CRS of Croatia: HTRS96/TM (EPSG: 3765) in the horizontal direction and HVRS71 (EPSG: 5610) in the vertical one (Figure 1).

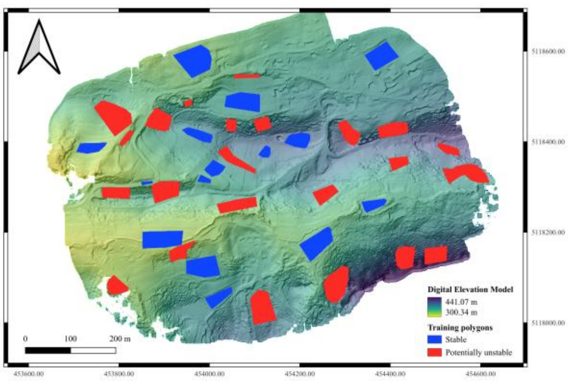

Figure 1. UAV LiDAR-derived terrain representation of the study area with training polygons

The obtained DEM was used as an input dataset to further perform geomorphometric calculations. A list of raster-based geomorphometric variables relevant to the problem of landslide susceptibility assessment in mountainous areas was defined on the basis of scientific literature, namely elevation, slope gradient, aspect, plan curvature, profile curvature, terrain position index, terrain roughness, flow accumulation, and the LS factor. All the raster layers were transformed to have the same spatial resolution, coordinate system, extents, and alignment to ensure that each cell had a set of input variables to be applied to the machine learning classifier.

Each layer containing raster variables was visually inspected in order to detect outliers and invalid values. If necessary, each raster layer was masked using the study area polygon to make sure that any non-spatial or outlying cell values did not affect model training. No-data values were not considered when selecting training data for machine learning analysis.

Training and testing reference polygons used for the purpose of supervised binary classification were created on the basis of geomorphological interpretation of high-resolution topography and additional engineering-geological maps. Two classes of terrain units, namely, potentially unstable and stable sites, were marked separately. The resulting reference dataset can be considered as an expert-based geomorphological interpretation supported by UAV LiDAR DEM and other data.

The entire processing workflow used in this work is designed to provide an opportunity for its easy reproduction. The input, output, and reference data, raster layers containing geomorphometric variables, results obtained from machine learning classifiers, and final figures were placed in separate folders to structure the dataset in a convenient manner. Each processing step, namely, geomorphometric feature generation and machine learning classification, was designed to be run with open-source geospatial software and Python.

3. Methodology

The proposed methodology represents an end-to-end, reproducible and open-source approach for terrain classification and susceptibility mapping using UAV LiDAR data. The workflow includes the following components: DEM processing, geomorphometry extraction, reference data preparation based on expert judgment, supervised machine learning classification, validation of the model and generation of susceptibility maps. All computational tasks have been performed in Jupyter Notebook using open-source Python libraries such as Rasterio, NumPy, pandas, SciPy, GeoPandas, Matplotlib and scikit-learn. The workflow automatically produces all processed rasters, tabulated data, results of the validation and susceptibility maps. Figure 2 shows full methodology process.

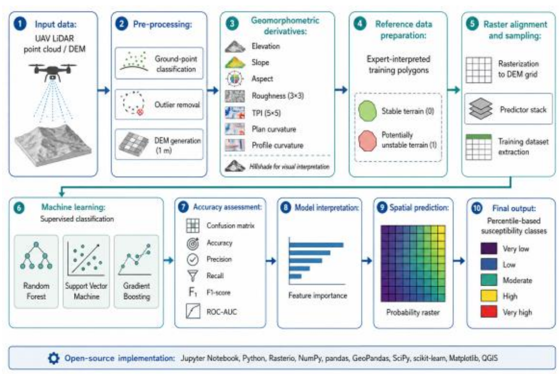

Figure 2. Reproducible open-source workflow used for UAV LiDAR-based terrain classification and relative hazard susceptibility mapping

3.1 Open-source processing workflow

The input dataset consists of a digital elevation model generated by a UAV LiDAR system, having a spatial resolution of 1 m. Processing was carried out in the Croatian coordinate reference system, defined by EPSG:3765 code. Before running the analyses, NoData pixels were converted to missing values. The workflow started with retrieving the metadata about the used DEM, i.e., the coordinate reference system, raster resolution, raster size, raster extent, number of cells, and elevation metrics. Afterward, the input DEM was used as a base for geomorphometric derivatives' calculations.

Intermediate and final results were saved in a directory structure with the processed raster derivatives, output rasters, table, and figures. This directory structure was used for ensuring transparency and traceability in each step of processing, ranging from the raw elevation data to the susceptibility model itself. All results could be reproduced using the same notebook by loading the original DEM and training polygons.

3.2 Geomorphometry feature extraction

Several morphometric, hydrologic, and surface roughness predictors were extracted from the DEM dataset collected using the UAV LiDAR system. In particular, the following parameters were used: elevation, slope, aspect, roughness, TPI, plan curvature, and profile curvature. To aid the visual analysis of the predictors, a hillshade layer was generated, but it was not used as an input feature for machine learning models since hillshade is a visualization tool generated based on the slope and aspect.

The slope was calculated using first-order elevations derivatives expressed in degrees. The aspect predictor was computed as the directional component of the terrain gradient ranging from 0° to 360°. The roughness of the terrain was computed based on a 3 × 3 moving window as the difference between the maximum and minimum elevation values within the moving window. It should be noted that this predictor was included in the model because landslides and potential instabilities are often associated with uneven microrelief, disturbed ground surface, and hummocky landforms. Topographic position index was computed based on a 5 × 5 moving window as the difference between the elevation value of a certain pixel and the average elevation value of the neighbouring pixels.

Plan and profile curvatures were computed as the second order derivatives of the DEM. Plan curvature reflects horizontal curvature and influences flow divergence and convergence. In turn, profile curvature reflects curvature in the downslope direction and determines the acceleration or deceleration of geomorphic processes along slope profiles. Curvature layers include extreme numerical artefacts that are masked to prevent any impact on the model due to unstable derivative values. All predictors were saved in GeoTIFF format and validated for consistency in dimensions, spatial reference systems, resolution, and grid alignment.

3.3 Reference data preparation

Training datasets for supervised classification were obtained based on the interpretation of expert training polygons that were created in QGIS. Training polygons were classified into two classes: stable terrain (class = 0) and potentially unstable terrain (class = 1). Interpreting polygons was based on a composite analysis of hillshades, slopes, roughness of terrain, TPI, and curvatures maps. Stable polygons were extracted from uniform geomorphic units without slope breaks, roughness, and curvature changes. Potentially unstable polygons were extracted from terrain units with features of steep slope, roughness, terrain discontinuity, and curvature breaks.

Vector polygons with information about their classes were imported with GeoPandas and validated against consistency in the CRS with DEM. Class distributions and polygon areas were calculated to characterize the training dataset. Polygons were then converted to raster with spatial parameters equal to DEM: same spatial extent, affine transformation, and pixel sizes. Obtained training data included two valid classes with classes equal to 0 for stable terrain and 1 for potentially unstable terrain.

A total of 63,202 pixels belonging to valid classes were extracted for supervised learning; among them, 25,846 pixels belonged to stable landforms, and 37,356 pixels belonged to potentially unstable landforms, which corresponds to 40.89% of stable landforms and 59.11% of potentially unstable landforms. Valid training pixels had all predictor raster values and valid classes assigned to them. In total, there were seven predictor variables and one target variable.

3.4 Machine-learning classification

The three machine learning algorithms evaluated were Random Forests, Support Vector Machines, and Gradient Boosting. They were chosen since they are popular and powerful models for classification tasks in geographic and terrain susceptibility modelling. Class balancing was applied in Random Forests, whereby an ensemble of decision trees was applied. In Support Vector Machines, a radial basis function kernel was used together with standardized input predictors from a scaling pipeline. Gradient Boosting was implemented as tree-based boosting algorithm with low learning rates and limited tree depths.

Stratified splitting was used to create two sets of data; training and testing set based on a ratio of 70:30. The number of samples in the training set was 44,241 whereas that in the testing set was 18,961. In both cases, the input predictor variables included: elevation, slope, aspect, terrain roughness, TPI, plan curvature and profile curvature.

3.5 Accuracy assessment and susceptibility mapping

Accuracy assessment was estimated by constructing the confusion matrices and computing such metrics as overall accuracy, precision, recall, F1-score, and ROC-AUC. This choice of metrics allowed to evaluate the overall classification accuracy and also to estimate the potential quality issues related to the potentially unstable class. Here precision is associated with the reliability of the predictions of the potentially unstable classes, recall indicates the ability of models to detect reference unstable pixels. F1-score combines these two features. Finally, ROC-AUC allows evaluating class separability in terms of the probability output.

For the tree-based algorithms – Random Forest and Gradient Boosting – feature importance was estimated to evaluate which geomorphometric variables had the highest contribution to classification. The SVM algorithm with a radial basis function kernel does not offer an out-of-the-box way of estimating

https://doi.org/10.5194/isprs-archives-L-4-W1-2026-127-2026

feature importance; therefore, it was tested only in terms of classification accuracy.

The selection of the best-performing model was done in terms of the F1-score and ROC-AUC values. Then this model was further applied to all non-null pixels of the full predictor stack to create a continuous probability raster showing the probability of potentially unstable terrain. Next, this raster was classified into susceptibility classes using the threshold values.

Two options of classification have been considered. In the first one, probabilities were set to fixed threshold values 0.2, 0.4, 0.6, and 0.8 in order to create five susceptibility classes – very low, low, moderate, high, and very high. Another option of classification involved setting threshold values to the specific percentiles of the probability distribution within the area. The 20th, 40th, 60th, and 80th percentiles were used to define five relative susceptibility classes: very low, low, moderate, high, and very high. In the current case, percentile values equalled 0.0867, 0.6267, 0.9233, and 0.9933. The latter approach was selected for visualization because it provided a better geographical balance between susceptibility classes.

4. Results

4.1 Geomorphometric characteristics of the UAV LiDAR DEM

The UAV LiDAR DEM generated a comprehensive representation of the topography in the study area. The elevations of the area varied from 300.03 m to 445.05 m, with a mean value of 357.31 m and a standard deviation of 25.09 m, revealing that the terrain surface was moderately heterogeneous, with adequate variations in elevation to carry out geomorphometric studies.

Slopes derived from the DEM showed great local variability, ranging from almost flat surfaces at 0.002° to highly inclined local slopes of 83.86°. The mean and median slopes were 19.33° and 15.50°, respectively, indicating the occurrence of moderate and steep terrain areas. Roughness, which was calculated using a 3 × 3 window approach, varied between 0.00 m to 17.39 m, with a mean of 1.03 m. It revealed that there existed a high degree of locally rough terrain surfaces, which might be associated with slope breaks, erosional surfaces, man-made terrains, and possibly unstable terrains.

TPI values ranged from -5.77 m to 10.68 m, encompassing the identification of local depressions, convex landforms, and relative differences in terrain positions. Plan curvature and profile curvature had averages of almost zero but revealed locally extreme values suggestive of variations in horizontal and vertical directions of the slopes. These predictors were retained as variables in modelling because they were relevant in determining local slopes, terrain convergences and divergences, and slope profiles of the terrain instability. Table 1 presents summarized statistics of DEM and predictors used.

Variable	Min.	Max.	Mean	Median	Std. dev.
Elevation (m)	300.03	445.05	357.31	356.39	25.09
Slope (°)	0.002	83.86	19.33	15.50	14.56
Aspect (°)	0.00	359.99	164.17	145.43	81.30
Roughness 3x3 (m)	0.00	17.39	1.03	0.76	0.97
TPI 5 x 5 (m)	-5.77	10.68	0.00008	-0.00070	0.28
Plan curvature	-9.99	9.99	-0.00005	-0.00027	0.64
Profile curvature	-2.09	4.44	0.00097	0.00029	0.12

Table 1. Summary statistics of the DEM and geomorphometry predictors used for machine-learning

4.2 Reference dataset and model performance

The expert-interpreted polygons were rasterized to the digital elevation model (DEM) grid to be used as reference data for supervised classification. The raster reference database included 63,202 pixels of which 25,846 belonged to stable landform and 37,356 pixels belonged to potentially unstable landform, which means that there were 40.89% stable and 59.11% potentially unstable reference pixels. The final machine learning database had seven predictor variables and one binary target class.

The database was divided into training and testing sets using a stratified split in proportion 70:30%. The training set included 44,241 samples, while the testing set included 18,961 samples. Three algorithms were tested, namely, Random Forest, Support Vector Machine, and Gradient Boosting. All of them exhibited high classification quality, and Random Forest demonstrated the best results.

Specifically, Random Forest had the following metrics: 0.9792 accuracy, 0.9789 precision, 0.9861 recall, 0.9825 F1-score, and 0.9976 ROC-AUC. The second algorithm was also quite successful with 0.9647 accuracy, 0.9716 precision, 0.9685 recall, 0.9701 F1-score, and 0.9937 ROC-AUC. The third one, however, scored lower but still very high numbers of 0.9377 accuracy, 0.9782 precision, 0.9150 recall, 0.9455 F1-score, and 0.9836 ROC-AUC.

The confusion matrix for Random Forest showed that 7,516 pixels of stable lands and 11,051 pixels of potentially unstable lands were correctly classified, while 238 stable lands were misclassified as potentially unstable lands and vice versa 156 potentially unstable lands were misclassified as stable lands. Summarized statistics is shown in Table 2 and Figure 3.

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
RF	0.9792	0.9789	0.9861	0.9825	0.9976
GB	0.9647	0.9716	0.9685	0.9701	0.9937
SVM	0.9377	0.9782	0.9150	0.9455	0.9836

Table 2. Model performance metrics for Random Forest (RF), Gradient Boosting (GB) and Support Vector Machine (SVM)

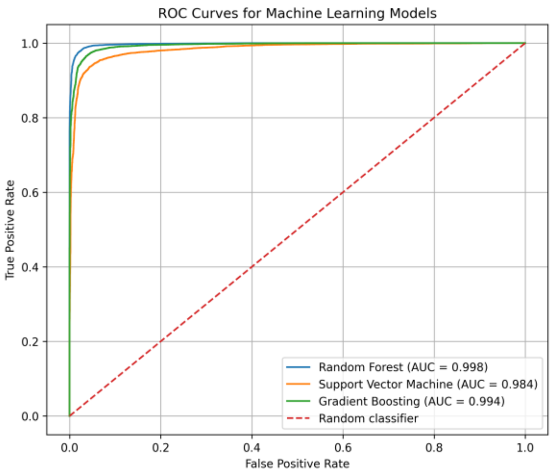

Figure 3. Model performance assessment based on the ROC curves for Random Forest, Support Vector Machine, and Gradient Boosting

4.3 Feature importance

In feature-importance analysis, the following predictors stood out among all of them for the chosen tree-based models: Random Forest (RF) and Gradient Boosting (GB). In the RF model, the most important feature turned out to be terrain roughness, with an importance level of 0.2933, with slope (0.2346) and elevation (0.2153) being next after roughness. Terrain position index (TPI), aspect, profile curvature, and plan curvature also had a certain influence on classification.

First of all, roughness was the most significant characteristic because local surface relief features play the key role in determining potential instability. It should be said that roughness is one of the expected properties of disturbance in high-resolution LiDAR-derived DEMs. The fact that slope turned out to be very important is evidence that terrain inclination plays an important role in instability caused by gravitation and erosion processes. As for elevation, the fact that its importance was relatively high means that unstable polygons in the study area are related to certain areas according to altitude.

Finally, when we turn to the gradient boosting model (GB), it shows us the strongest importance of roughness (0.6665), with elevation taking the second place (0.1879) and TPI, aspect, profile curvature, slope, and plan curvature having a lesser effect.

4.4 Hazard susceptibility mapping

The Random Forest model was chosen for spatial prediction based on its maximum performance in terms of F1-score and ROC-AUC score compared to other models. The trained model was then utilized for predicting spatial variability across all 517,237 valid pixels of the predictor stack, generating a continuous raster layer that indicated the probability of the land being prone to landslides.

Using four pre-defined fixed probability thresholds at 0.2, 0.4, 0.6, and 0.8, 51.27% of the study area was categorized into the very high susceptibility category, signifying a significant shift towards higher predicted probabilities in the distribution of probabilities. Thus, a percentile approach was considered to facilitate relative interpretation of susceptibility within the study area. In particular, the thresholds of predicted probability values were defined using the 20th, 40th, 60th, and 80th percentiles, resulting in threshold values of 0.0867, 0.6267, 0.9233, and 0.9933, respectively (Figure 4).

The percentile approach provided a better distinction of the study area based on susceptibility level. High susceptibility zones were primarily located along extended slope break segments, morphologically complex terrain bands, and rough terrain zones. On the contrary, low susceptibility zones mainly covered relatively flat and smooth terrain surfaces. The generated susceptibility map is thus meant to be interpreted relatively where the highest susceptibility zone includes the top 20% of predicted probability values in the study area.

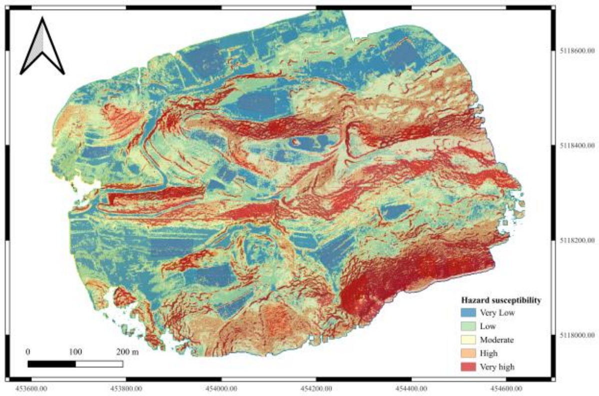

Figure 4. Percentile-based relative hazard susceptibility map derived from the Random Forest probability raster

5. Discussion

The results demonstrate that UAV LiDAR-based geomorphometric variables can serve as a reliable basis for high-resolution terrain classification and susceptibility mapping using supervised machine-learning algorithms. Using a DEM with a resolution of 1 m enabled the detailed depiction of terrain morphology and provided the possibility of extracting a number of local terrain properties, namely, slope, roughness, TPI, and curvature. These metrics have been shown to represent different aspects of terrain morphology like inclination, microtopography, relative terrain position, and morphological transitions of slope. Together they provided more information about terrain types compared to relying on one sole variable.

When comparing the performance of different machine-learning models, Random Forest showed the best results, with accuracy of 0.9792, F1 score of 0.9825, and ROC AUC of 0.9976. Gradient Boosting also exhibited outstanding performance, while SVM had slightly decreased recall for potentially unstable terrain types. The success of decision tree-based machine learning models is explained by their ability to model nonlinear interactions between geomorphometric input variables and target classes. In the current application, Random Forest provided the best trade-off between classifying potentially unstable terrain and avoiding false-positive classification of stable terrain, making it suitable for creating the susceptibility probability map.

According to the feature-importance analysis, terrain roughness served as the leading predictor in both decision tree-based models. The Random Forest was followed by slope and elevation, whereas in the case of Gradient Boosting, roughness became even more important. Such results suggest that expert-interpreted potentially unstable terrain polygons possess the characteristic of being locally irregular and having complex microtopography. This is geomorphologically reasonable since potentially unstable and disturbed slopes are usually associated with hummocky relief, small scarps, erosion traces, displaced debris, or locally irregular landforms. Besides, slope had a high importance value, indicating its central role as a condition for gravitational and erosion processes. The relatively high importance of elevation might be attributed to the fact that potentially unstable polygons are located in certain topographical positions and not spread across all elevations.

Curvature variables demonstrated less importance than slope and roughness, but they still participated in the classification process. The relatively low importance of curvatures can be caused by the fact that they are highly sensitive to noise and local micro-topographic variations, especially at 1 m resolution. Plan and profile curvatures should theoretically assist in recognizing converging and diverging terrain forms, local slope breaks, and convex/concave parts of a slope, but they play their role dependent on the scale of computation and local terrain characteristics. In future studies, several scales of curvature calculations or its preliminary smoothing should be considered.

The results of the susceptibility-mapping experiment show that applying fixed probability thresholds resulted in a significant part of very high susceptibility terrain. Such a situation means that the values calculated by Random Forest have a skewed distribution over the territory. Moreover, potentially unstable and stable polygons have morphological similarities; thus, there is a significant proportion of pixels with morphologically similar conditions as potentially unstable polygons. Therefore, percentile-based susceptibility classification was applied for another way of susceptibility evaluation. The percentile-based classification provided better discrimination between different classes and facilitated visual analysis of potentially unstable terrain zones. However, such susceptibility map should be interpreted as a ranking based on a relative probability of being unstable terrain within the territory under study.

One of the advantages of the current research is its full openness of the analysis workflow implemented using open-source geospatial software and libraries. The presented workflow includes several major steps like DEM import, geomorphometric derivatives computing, reference polygons rasterization, supervised machine learning training, accuracy assessment, feature importance analysis, and susceptibility map creation. All these steps create automatically stored results in raster, table, and figure format, allowing for easy verification and reproducibility. In the scope of FOSS4G, apart from the final product, openness and transferability of the developed workflow is especially important for assessing its practical value.

Still, there are several limitations that should be considered while interpreting the results. Firstly, the reference polygons are created via experts' interpretation of UAV LiDAR-based terrain morphology rather than an inventory based on the actual landslide presence. Thus, the map should be understood as a "potentially unstable terrain" rather than the actual landslide map. Secondly, the validation was done via random stratified train-test split of raster pixels, which does not account for potential spatial autocorrelation between neighbours. Spatial block validation should be used in order to estimate how the model works on other parts of the study area. Thirdly, in the presence of anthropogenic objects, roads, and terraces with high roughness values, there is a risk of getting false-positive results due to their similarity with unstable training polygons.

All things considered; the results confirm the applicability of UAV LiDAR-based geomorphometric analysis coupled with machine learning for creating a detailed terrain susceptibility map. The provided workflow can be used as a reproducible tool for identifying territories with possible unstable landforms for further field inspections or engineering geological analysis. Some directions for further improvement include adding more reference polygons, validation via fieldwork, spatial cross-validation, considering multi-scale geomorphometric derivatives, and incorporating additional environmental predictors like flow accumulation and LS factors.

6. Conclusion

This paper presents an open-source workflow for producing a reliable and reproducible high-resolution classification of the terrain units based on relative hazard susceptibility mapping using UAV LiDAR derived geomorphometric features and supervised machine learning approaches. The workflow includes DEM processing, terrain metrics calculation, creation of reference polygons by an expert, model training, accuracy assessment, feature importance analysis, and generation of susceptibility maps all carried out in an open-source Python/Jupyter notebook environment.

It has been demonstrated that UAV LiDAR derived geomorphometry is capable of providing a good basis for terrain units separation, which could be potentially unstable. Out of three tested models, Random Forest proved to be the most accurate, delivering the highest values of the classification accuracy (0.9792), F1-score (0.9825), and ROC-AUC (0.9976). Gradient Boosting and SVM provided very high accuracy too, however, Random Forest proved to yield more balanced results. In terms of feature importance, it appeared that three features contributed the most to the final prediction: terrain roughness, slope and elevation.

The final relative susceptibility map was generated based on probabilities obtained during Random Forest classification and reclassification according to percentile thresholds. This approach allowed to produce a classification map of the relative susceptibility to landslides better than the fixed threshold classification of the probabilities. In general, the areas characterized as having the highest susceptibility were mostly morphologically heterogeneous, characterized by slope breaks and rugged surfaces, while smoother homogeneous regions were classified as less susceptible.

The main strength of this paper resides not in the production of a susceptibility map, but in the presentation of the methodology, which can be transferred to other areas. The whole pipeline of data processing can be reproduced with any DEM and expert-provided polygons. Thus, future work should consider the integration of the field validated landslides inventory into the training set, application of block-based testing, use of various scales in the calculation of geomorphological features, and integration of hydrological features.

Acknowledgements

This paper was supported by the 2026 Call for Proposals for Young Researchers at the University North.
The complete reproducible workflow is available on: github.com/nkranjcic/foss4g-2026-uav-lidar-terrain-ml (29 July 2026).
Data used for this research are collected as a part of the research project funded by Croatian Science Foundation under Grant

HRZZ-IP-2022-10-1296 - Seismic ground motion amplification induced by topographic irregularity in Northern Croatia – SIGMATOPCRO.

References

Al-Najjar, H.A.H., Pradhan, B., Kalantar, B., Sameen, M.I., Santosh, M., Alamri, A., 2021. Landslide Susceptibility Modeling: An Integrated Novel Method Based on Machine Learning Feature Transformation. *Remote Sensing* 13, 3281. doi.org/10.3390/rs13163281

Barbarella, M., Di Benedetto, A., Fiani, M., 2021. Application of Supervised Machine Learning Technique on LiDAR Data for Monitoring Coastal Land Evolution. *Remote Sensing* 13, 4782. doi.org/10.3390/rs13234782

Bartmiński, P., Siłuch, M., Kociuba, W., 2023. The Effectiveness of a UAV-Based LiDAR Survey to Develop Digital Terrain Models and Topographic Texture Analyses. *Sensors* 23, 6415. doi.org/10.3390/s23146415

Butler, H., Chambers, B., Hartzell, P., Glennie, C., 2021. PDAL: An open source library for the processing and analysis of point clouds. *Computers & Geosciences* 148, 104680. doi.org/10.1016/j.cageo.2020.104680

Forkan, A., Both, A., Bellman, C., Duckham, M., Anderson, H., Radosevic, N., 2023. K-span: Open and reproducible spatial analytics using scientific workflows. *Front. Earth Sci*. 11, 1130262. doi.org/10.3389/feart.2023.1130262

Kavzoglu, T., Teke, A., 2022. Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost). *Arab J Sci Eng* 47, 7367–7385. doi.org/10.1007/s13369-022-06560-8

Maesen, P., Salingros, E., 2024. Introduction to Reproducible Geospatial Analysis and Figures in R: A Tutorial Article. *Data* 9, 58. doi.org/10.3390/data9040058

Scheiber, L., Zühlsdorff, V., Nong, D.H., Ngo, T.S., Downes, N.K., Bachofer, F., Nguyen, H.Q., Garschagen, M., Reimuth, A., 2026. Monitoring urban green space for climate-resilient development in the face of rapid urbanization: A tale of two Vietnamese cities. *Remote Sensing Applications: Society and Environment* 41, 101820. doi.org/10.1016/j.rsase.2025.101820

Storch, M., De Lange, N., Jarmer, T., Waske, B., 2023. Detecting Historical Terrain Anomalies With UAV-LiDAR Data Using Spline-Approximation and Support Vector Machines. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing* 16, 3158–3173. doi.org/10.1109/JSTARS.2023.3259200

Storch, M., Kisliuk, B., Jarmer, T., Waske, B., de Lange, N., 2025. Comparative analysis of UAV-based LiDAR and photogrammetric systems for the detection of terrain anomalies in a historical conflict landscape. *Science of Remote Sensing* 11, 100191. doi.org/10.1016/j.srs.2024.100191

Xu, Q., Yordanov, V., Amici, L., Brovelli, M.A., 2024. Landslide susceptibility mapping using ensemble machine learning methods: a case study in Lombardy, Northern Italy. *International Journal of Digital Earth* 17, 2346263. doi.org/10.1080/17538947.2024.2346263